Appendix.—Accuracy of the Data

INTRODUCTION

The data presented in this publication are based on the 1980 census sample. The data are estimates of the actual figures that would have resulted from a complete count. Estimates can be expected to vary from the complete count result, because they are subject to two basic types of error—sampling and nonsampling. The sampling error in the data arises from the selection of persons and housing units to be included in the sample. The nonsampling error, which affects both sample and complete count data, is the result of all other errors that may occur during the collection and processing phases of the census. A more detailed discussion of both sampling and nonsampling error and a description of the estimation procedure are given in this appendix.

SAMPLE DESIGN

While every person and housing unit in the United States was enumerated on a questionnaire that requested certain basic demographic information (e.g., age, race, relationship), a sample of persons and housing units was enumerated on a questionnaire that requested additional information. The basic sampling unit for the 1980 census was the housing unit, including all occupants. For persons living in group quarters, the sampling unit was the person. Two sampling rates were employed. In counties, incorporated places, and minor civil divisions estimated to have fewer than 2,500 persons (based on precensus estimates), one-half of all housing units and persons in group quarters were to be included in the sample. In all other places, one-sixth of the housing units or persons in group quarters were sampled. The purpose of this scheme was to provide relatively more reliable estimates for small places. When both sampling rates were taken into account across the Nation, approximately 19 percent of the Nation's housing units were included in the census sample.

The sample designation method depended on the data collection procedures. In about 95 percent of the country, the census was taken by the mailout/mailback procedure. For these areas, the Bureau of the Census either purchased a commercial mailing list which was updated and corrected by Census Bureau field staff, or prepared a mailing list by canvassing and listing each address in the area prior to Census Day. These lists were computerized, and every sixth unit (for 1-in-6 areas) or every second unit (for 1-in-2 areas) was designated as a sample unit by computer. Both of these lists were also corrected by the Post Office.

In non-mailout/mailback areas, a blank listing book with designated sample lines (every sixth or every second line) was prepared for the enumerator. Beginning about Census Day, the enumerator systematically canvassed the area and listed all housing units in the listing book in the order they were encountered. Completed questionnaires, including sample information for any housing unit which was listed on a designated sample line, were collected.

In both types of data collection procedure areas, an enumerator was responsible for a small geographic area known as an enumeration district, or ED. An ED usually represented the average workload area for one enumerator.

ERRORS IN THE DATA

Since the data in this publication are based on a sample, they may differ somewhat from complete count figures that would have been obtained if all housing units, persons within those housing units, and persons living in group quarters had been enumerated using the same questionnaires, instructions, enumerators, etc. The deviation of a sample estimate from the average of all possible samples is called the sampling error. The standard error of a survey estimate is a measure of the variation among the estimates from the possible samples and thus is a measure of the precision with which an estimate from a particular sample approximates the average result of all possible samples. The sample estimate and its estimated standard error permit the construction of interval estimates with prescribed confidence that the interval includes the average result of all possible samples. The method of calculating standard errors and confidence intervals for the data in this report is given below.

In addition to the variability which arises from the sampling procedures, both sample data and complete-count data are subject to nonsampling error. Nonsampling error may be introduced during each of the many extensive and complex operations used to collect and process census data. For example, operations such as editing, reviewing, or handling questionnaires may introduce error into

the data. A more detailed discussion of the sources of nonsampling error is given in the section on "Control of Nonsampling Errors" in this appendix.

Nonsampling error may affect the data in two ways. Errors that are introduced randomly will increase the variability of the data and should therefore be reflected in the standard error. Errors that tend to be consistent in one direction will make both sample and complete-count data biased in that direction. For example, if respondents consistently tend to underreport their income, then the resulting counts of households or families by income category will be skewed toward the lower income categories. Such biases are not reflected in the standard error.

Calculation of Standard Errors

Totals and Percentages—Tables J and K in this appendix contains the information necessary to calculate the standard errors of all census sample estimates in this report. The standard errors of the Current Population Survey (CPS) estimates are given along with the CPS estimates in the text of this report. The CPS standard errors do not require any further adjustments and may be used as they appear. In order to calculate standard errors and census sample estimates, the steps in this section must be followed. To perform the calculations of census standard errors, it is necessary to know the unadjusted standard error for the characteristic, given in tables J or K, that would result under a simple random sample design (of persons) and estimation technique; the adjustment factor for the particular characteristic estimated, is given in table L. The adjustment factors reflect the effects of the actual sample design and complex ratio estimation procedure used for the 1980 census.

To calculate the approximate standard error of a census estimate, follow the steps given below:

a. Obtain the unadjusted standard error from table J or K (or from the formula given below the table) for the estimated total or percentage, respectively.

b. Use table L to obtain the factor for the Ancestry characteristic. Multiply the unadjusted standard error by this factor. If the estimate is a cross-tabulation of more than one characteristic, use the largest factor.

As is evident from the formulas below tables J and K, the unadjusted standard errors of zero estimates or of very small estimated totals or percentages approach zero. This is also the case for very large percentages or estimated totals that are close to the size of the tabulation areas to which they correspond. These estimated totals and percentages are, nevertheless, still subject to sampling and nonsampling variability, and an estimated standard error of zero (or a very small standard error) is not appropriate.

For estimated percentages that are less than 2 or greater than 98, use the *unadjusted* standard errors in table K that appear in the "2" or "98" row. For an estimated total that is less than 50 or within 50 of the total size of the tabulation area, use an *unadjusted* standard error of 16.

An illustration of the use of the tables is given in a later section of this appendix.

Differences—The standard errors estimated from these tables are not directly applicable to differences between two sample estimates. In order to estimate the standard error of a difference, the tables are to be used somewhat diffrently in the following three situations:

a. For the difference between a sample estimate and a complete-count value, use the standard error of the sample estimate.

b. For the difference between (or sum of) two census sample estimates, the appropriate standard error is approximately the square root of the sum of the two individual standard errors squared; that is, for standard errors Se_x and Se_y of estimates x and y:

$$Se_{(x+y)} = Se_{(x-y)} = \sqrt{(Se_x)^2 + (Se_y)^2}$$

This method, however, will underestimate (overestimate) the standard error if the two items in a sum are highly positively (negatively) correlated or if the two items in a difference are highly negatively (positively) correlated. This method may also be used for the difference between (or sum of) sample estimates from two censuses or between a census sample and another survey such as the CPS. The standard error for estimates not based on the 1980 census sample and not given in this report, must be obtained from an appropriate source outside of this publication.

c. For the difference between two census sample estimates, one of which is a subclass of the other, use the tables directly where the calculated difference is the estimate of interest.

Confidence Intervals

A sample estimate and its estimated standard error may be used to construct confidence intervals about the estimate. These intervals are ranges that will contain the average value of the estimated characteristic that results over all possible samples, with a known probability. For example, if all possible samples that could result under the 1980 census sample design were independently selected and surveyed under the same condition and if the estimate and its estimated standard error were calculated for each of these samples, then:

(1) Approximately 68 percent of the intervals from one estimated standard error below the estimate to one estimated standard error above the estimate would contain the average result from all possible samples; and

(2) Approximately 95 percent of the intervals from two estimated standard errors below the estimate to two estimated standard errors above the estimate would contain the average result from all possible samples.

The intervals are referred to as 68 percent and 95 percent confidence intervals, respectively.

The average value of the estimated characteristic that could be derived from all possible samples is or is not contained in any particular computed interval. Thus we cannot make the statement that the average value has a certain probability of falling between the limits of the calculated confidence interval. Rather, one can say with a specified probability or confidence that the calculated confidenc

interval includes the average estimate from all possible samples (approximately the complete-count value).

Confidence intervals may also be constructed for the difference between two sample figures. This is done by computing the difference between these figures, obtaining the standard error of the difference (using the formula given earlier) and then forming a confidence interval for this estimated difference as above. One can then say with specified confidence that this interval includes the difference that would have been obtained by averaging the results from all possible samples.

The estimated standard errors given in this report do not include all portions of the variability due to nonsampling error that may be present in the data. Thus, the standard errors calculated represent a lower bound of the total error. As a result, confidence intervals formed using these estimated standard errors may not meet the stated levels of confidence (i.e., 68 or 95 percent). Thus, some care must be exercised in the interpretation of the data in this publication based on the estimated standard errors.

For more information on confidence intervals and nonsampling error see any standard sampling theory text.

Use of Tables to Compute Standard Errors

Table 3a shows that for the State of Oklahoma, out of 3,025,290 persons, 400,283 were reported to be of English ancestry (single). The procedure for obtaining the standard error of 400,283 will be demonstrated.

The unadjusted standard error for the estimated total is obtained from table J or from the formula below table J. In order to avoid interpolation, the use of the formula will be demonstrated here. By the formula, the unadjusted standard error, Se, is:

$$Se = \sqrt{5\,(400{,}283)\left(1-\frac{400{,}283}{3{,}025{,}290}\right)}$$

$$= 1{,}318 \quad \text{persons.}$$

The standard error of the estimated 400,283 persons of English ancestry in Oklahoma is found by multiplying the unadjusted standard error, 1,318, by the appropriate adjustment factor Table L lists the adjustment factor for the characteristic "European (excluding Spaniard) Single Ancestry." It is shown to be 1.5. Thus, the estimated standard error is 1,318 x 1.5 or 1,977.

The estimated percent of persons from English ancestry (single) in Oklahoma is 13.2. From the formula shown in table K, the unadjusted standard error is found to be .04. Thus the standard error for the estimated percent of persons with English ancestry in Oklahoma is 1.5 x .04 = .06.

A note of caution concerning numerical values is necessary. Standard errors of percentages derived from table K are approximate. Calculations can be expressed to several decimal places, but to do so would indicate more precision in the data than is justifiable. Final results should contain no more than one decimal place when the standard error is one percentage point (i.e., 1.0) or more.

In the previous example, the standard error of the estimated 400,283 persons of English ancestry (single) in Oklahoma is found to be 1,977. Thus, a 95-percent confidence interval for this estimated total is found to be:

400,283 − 2(1,977) to 400,283 + 2(1,977)

or

396,329 to 404,237.

One can say with about 95-percent confidence that this interval includes the actual value that would have been obtained by averaging the results from all possible samples.

The calculation of standard errors and confidence intervals will be illustrated when a difference of two sample estimates is obtained. For example, out of 23,667,902 persons in California 1,827,247 have English ancestry (single). Thus, the percentage of persons with English ancestry (single) in California is 7.7 percent. The unadjusted standard error from the formula in table K is .01 percent. From table L the adjustment factor is found to be 1.5 for "European (excluding Spaniard) Single Ancestry." Thus, the approximate standard error of the percentage (7.7 percent) is .01 x 1.5 = .02.

Suppose that one wishes to obtain the standard error of the difference between Oklahoma and California of the percentages of persons with English ancestry (single). The difference in the percentages of interest for the two States is:

13.2 − 7.7 = 5.5 percent.

Using the results of the previous example:

$$Se\,(5.5) = \sqrt{(Se(13.2))^2 + (Se(7.7))^2}$$

$$= \sqrt{(.06)^2 + (.02)^2}$$

$$= .06 \text{ percent.}$$

The 95-percent confidence interval for the difference is formed as before:

5.5 − 2 (.06) to 5.5 + 2 (.06)

or

5.4 to 5.6.

One can say with 95-percent confidence that the interval includes the actual difference that would have been obtained by averaging the results from all possible samples.

ESTIMATION PROCEDURE

The estimates which appear in this publication were obtained from an iterative ratio estimation procedure which resulted in the assignment of a weight to each sample person. For any given tabulation area, a characteristic total was estimated by summing the weights assigned to the persons in the tabulation area which possessed the characteristic. Estimates of family characteristics were based on the weights assigned to the family members designated as householders. Each sample person was assigned exactly one weight to be used to produce estimates of all characteristics. For example, if the weight given to a sample person had the value five, all characteristics of that person or housing unit would be tabulated with a weight of five. The estimation procedure, however, did assign weights which vary from person to person.

The estimation procedure used to assign the weights was performed in geographically defined "weighting areas." Weighting areas were generally formed of adjoining portions of geography, which closely agreed with census tabulation areas within counties. Weighting areas were required to have a minimum sample of 400 persons. Weighting areas were never allowed to cross state or county boundaries. In small counties with a

sample count of less than 400 persons, the minimum required sample condition was relaxed to permit the entire county to become a weighting area.

Within a weighting area, the ratio estimation procedure for persons was performed in three stages. For persons the first stage employed 17 household type groups. The second stage used two groups: householders and non-householders. The third stage could potentially use 160 age-sex-race-Spanish origin groups. The stages were as follows:

Stage I—Type of Household

Group	*Persons in Housing Units With a Family With Own Children Under 18.*
1	2 persons in housing unit
2	3 persons in housing unit
3	4 persons in housing unit
4	5 to 7 persons in housing unit
5	8 or more persons in housing unit
	Persons in Housing Units With a Family Without Own Children Under 18.
6-10	2 persons in housing unit through 8 or more persons in housing unit
	Persons in All Other Housing Units
11	1 person in housing unit
12-16	2 persons in housing unit through 8 or more persons in housing unit
17	*Persons in Group Quarters*

Stage II—Householder/Nonhouseholder

Group	
1	Householder
2	Non-householder (including persons in group quarters)

Stage III—Age/Sex/Race/Spanish Origin

Group	
	White Race
	Persons of Spanish Origin
	Male
1	0 to 4 years of age
2	5 to 14 years of age
3	15 to 19 years of age
4	20 to 24 years of age
5	25 to 34 years of age
6	35 to 44 years of age
7	45 to 64 years of age
8	65 years of age or older
	Female
9-16	Same age categories as groups 1 to 8
	Persons Not of Spanish Origin
17-32	Same age and sex categories as groups 1 to 16
	Black Race
33-64	Same age-sex-Spanish origin categories as groups 1 to 32
	Asian, Pacific Islander Race
65-96	Same age-sex-Spanish origin categories as groups 1 to 32
	Indian (American) or Eskimo or Aleut Race
97-128	Same age-sex-Spanish origin categories as groups 1 to 32
	Other Race (includes those races not listed above)
129-160	Same age-sex-Spanish origin categories as groups 1 to 32

Within a weighting area, the first step in the estimation procedure was to assign each sample person record an initial weight. This weight was approximately equal to the inverse of the probability of selecting a person for the census sample.

The next step in estimation procedure was to combine, if necessary, the groups in each of the three stages prior to the repeated ratio estimation in order to increase the reliability of the ratio estimation procedure. For the first and second stages, any group that did not meet certain criteria concerning the unweighted sample count or the ratio of the complete count to the initially weighted sample count was combined or collapsed with another group in the same stage according to a specified collapsing pattern. At the third stage, the "other" race category was collapsed with the "White" race category before the above collapsing criteria as well as an additional criterion concerning the number of complete-count persons in each category were applied.

As the final step, the initial weights underwent three stages of ratio adjustment which used the groups listed above. At the first stage, the ratio of the complete census count to the sum of the initial weights for each sample person was computed for each stage I group. The initial weight assigned to each person in a group was then multiplied by the stage I group ratio to produce an adjusted weight. In stage II, the stage I adjusted weights were again adjusted by the ratio of the complete census count to the sum of the stage I weights for sample persons in each stage II group. Finally, the stage II weights were adjusted at stage III by the ratio of the complete census count and the sum of the stage II weights for sample persons in each stage III group. The three stages of adjustment were performed twice (two iterations) in the order given above. The weights obtained from the second iteration for stage III were assigned to the sample person records. However, to avoid complications in rounding for tabulated data, only whole number weights were assigned. For example, if the final weight for the persons in a particular group was 7.2, then one-fifth of the sample persons in this group were randomly assigned a weight of 8 and the remaining four-fifths received a weight of 7.

The estimates produced by this procedure realize some of the gains in sampling efficiency that would have resulted if the population had been stratified into the ratio estimation groups before sampling, and the sampling rate had been applied independently to each group. The net effect is a reduction in both the standard error and the possible bias of most estimated characteristics to levels below what would have resulted from simply using the initial (unadjusted) weight. A by-product of this estimation procedure is that the estimates from the sample will, for the most part, be consistent with the complete-count figures for the population and housing unit groups used in the estimation procedure.

CONTROL OF NONSAMPLING ERROR

As mentioned above, nonsampling error is present in both sample and complete-count data. If left unchecked, this error could introduce serious bias into the data, the variability of which could increase dramatically over that which would result

purely from sampling. While it is impossible to completely eliminate nonsampling error from an operation as large and complex as the 1980 census, the Bureau of the Census attempted to control the sources of such error during the collection and processing operations. The primary sources of nonsampling error and the programs instituted for control of this error are described below. The success of these programs, however, was contingent upon how well the instructions were actually carried out during the census. To the extent possible, both the effects of these programs and the amount of error remaining after their application will be evaluated.

Undercoverage—It is possible for some housing units or persons to be entirely missed by the census. This undercoverage of persons and housing units can introduce biases into the data. Several extensive programs that were developed to focus on this important problem are explained below.

- The Postal Service reviewed mailing lists and reported housing unit addresses which were missing, undeliverable, or duplicated in the listings.
- The purchased commercial mailing list was updated and corrected by a complete field review of the list of housing units during a precanvass operation.
- A record check was performed to reduce the undercoverage of individual persons in selected areas. Independent lists of persons, such as driver's license holders, were matched with the household rosters in the census listings. Persons not matched to the census rosters were followed up and added to the census counts if they were found to have been missed.
- A recheck of units initially classified as vacant or nonexistent was utilized to further reduce the undercoverage of persons.

More extensive discussions of programs developed to reduce undercoverage will be published as the analyses of those programs are completed.

Respondent and Enumeration Error—The person answering the questionnaire or responding to the questions posed by an enumerator could serve as a source of error by offering incorrect or incomplete information. To reduce this source of error, questions were phrased as clearly as possible based on precensus tests and detailed instructions for completing the questionnaire were provided to each household. In addition, respondents' answers were edited for completeness and consistency and followed up as necessary. For example, if labor force items were incomplete for a person 15 years or older, long-form field edit procedures would recognize the situation and a followup attempt to obtain the information would be made.

The enumerator may misinterpret or otherwise incorrectly record information given by a respondent; may fail to collect some of the information for a person or household; or may collect data for households that were not designated as part of the sample. To control these problems, the work of enumerators was carefully monitored. Field staff were prepared for their tasks by using standardized training packages which included experience in using census materials. A sample of the households interviewed by enumerators for nonresponse was reinterviewed to control for the possibility of data for fabricated persons being submitted by enumerators. Also, the estimation procedure was designed to control for biases that would result from the collection of data from households not designated for the sample.

Processing Error—The many phases of processing the census represent potential sources for the introduction of nonsampling error. The processing of the census questionnaires includes the field editing, followup, and transmittal of completed questionnaires; the manual coding of write-in responses; and the electronic data processing. The various field, coding and computer operations undergo a number of quality control checks to insure their accurate application.

Nonresponse—Nonresponse to particular questions on the census questionnaire allows for the introduction of bias into the data since the characteristics of the nonrespondents have not been observed, and may differ from those reported by respondents. As a result, any allocation procedure using respondent data may not completely reflect this difference either at the element level (individual person or housing unit) or on the average. Some protection against the introduction of large biases is afforded by minimizing nonresponse. In the census, nonresponse was substantially reduced during the field operations by the various edit and followup operations aimed at obtaining a response for every question. Characteristics of the nonrespondents remaining after this operation were allocated by computer using reported data for a person or housing unit with similar characteristics. The allocation procedure is described in more detail below.

EDITING OF UNACCEPTABLE DATA

The objective of the processing operation is to produce a set of statistics that describes the population as accurately and clearly as possible. To meet this objective, certain unacceptable entries were edited.

In the field, questionnaires were reviewed for omissions and certain inconsistencies by a census clerk or an enumerator and, if necessary, a followup was made to obtain missing information. In addition, a similar review of questionnaires was done in the central processing offices. As a rule, however, editing was performed by hand only when it could not be done effectively by machine.

As one of the first steps in editing, the configuration of marks on the questionnaire column was scanned electronically to determine whether it contained information for a person or a housing unit or merely spurious marks. If the column contained entries for at least two of the basic characteristics (relationship, sex, race, age, marital status, Spanish origin), the inference was made that the marks represented a person. In cases in which two or more basic characteristics were available for only a portion of the people in the unit, other information on the questionnaire provided by an enumerator was used to determine the total number of persons. Names were not used as a criterion of the presence of a person because the electronic scanning did not distinguish any entry in the name space.

If any characteristic for a person or a housing unit was still missing when the questionnaires reached the central processing offices, they were supplied by allo-

cation. Allocations, or assignments of acceptable codes in place of unacceptable entries, were needed most often when there was no entry for a given item or when the information reported for a person or housing unit on that item was inconsistent with other information for the person or housing unit. As in previous censuses, the general procedure for changing unacceptable entries was to assign an entry for a person or housing unit that was consistent with entries for other persons or units with similar characteristics. Thus, a person who was reported as a 20-year-old son of the householder, but for whom marital status was not reported, was assigned the same marital status as that of the last one processed in the same age group. The assignment of acceptable codes in place of blanks or unacceptable entries enhances the usefulness of the data.

The editing process also includes another type of correction; namely, the assignment of a full set of characteristics for a person or a housing unit. When there was indication that a housing unit was occupied but the questionnaire contained no information for all or most of the people, although persons were known to be present, or when there was no information on the housing unit, a previously processed household was selected as a substitute, and the full set of characteristics for each substitute person or a housing unit was duplicated. These duplications fall into two classes: (1) "substitution for mechanical failure," e.g., when the questionnaire page was not properly microfilmed, and (2) "substitution for noninterview," e.g., when a housing unit was indicated as occupied but the occupants or housing unit characteristics were not listed on the questionnaire.

Specific tolerances were established for the number of computer allocations and substitutions that would be permitted. If the number of corrections was beyond tolerance, the questionnaires in which the errors occurred were clerically reviewed. If it was found that the errors resulted from damaged questionnaires, from improper microfilming, from faulty reading by FOSDIC of undamaged questionnaires, or from other types of machine failure, the questionnaires were reprocessed.

Table J. Unadjusted Standard Errors for Estimated Totals

[Based on a 1-in-6 simple random sample]

Estimated Total 1/	Size of publication area 2/								
	50 000	100 000	250 000	500 000	1 000 000	5 000 000	10 000 000	25 000 000	United States
50	16	16	16	16	16	16	16	16	16
100	22	22	22	22	22	22	22	22	22
250	35	35	35	35	35	35	35	35	35
500	50	50	50	50	50	50	50	50	50
1 000	70	70	70	70	70	70	70	70	70
2 500	110	110	110	110	110	110	110	110	110
5 000	150	150	160	160	160	160	160	160	160
10 000	200	210	220	220	220	220	220	220	220
15 000	230	250	270	270	270	270	270	270	270
25 000	250	310	340	350	350	350	350	350	350
75 000	-	310	510	570	590	610	610	610	610
100 000	-	-	550	630	670	700	710	710	710
250 000	-	-	-	790	970	1 090	1 100	1 100	1 120
500 000	-	-	-	-	1 120	1 500	1 540	1 570	1 580
1 000 000	-	-	-	-	-	2 000	2 120	2 190	2 230
5 000 000	-	-	-	-	-	-	3 540	4 470	4 940
10 000 000	-	-	-	-	-	-	-	5 480	6 910

1/ For estimated totals larger than 10 000 000, the standard error is somewhat larger than the table values. The formula given below should be used to calculate the standard error.

$$Se\ (\hat{Y}) = \sqrt{5\hat{Y}(1 - \frac{\hat{Y}}{N})}$$

N = Size of area

$\hat{Y}$ = Estimate of characteristic total

2/ Total count of persons in area if the estimated total is a person characteristic, or the total count of housing units in area if the estimated total is a housing unit characteristic.

Table K. Unadjusted Standard Error in Percentage Points for Estimated Percentages

[Based on a 1-in-6 simple random sample]

Estimated Percentage	Base of percentage 1/											
	500	750	1 000	1 500	2 500	5 000	7 500	10 000	25 000	50 000	100 000	250 000	500 000
2 or 98	1.4	1.1	1.0	0.8	0.6	0.4	0.4	0.3	0.2	0.1	0.1	0.1	0.1
5 or 95	2.2	1.8	1.5	1.3	1.0	0.7	0.6	0.5	0.3	0.2	0.2	0.1	0.1
10 or 90	3.0	2.4	2.1	1.7	1.3	0.9	0.8	0.7	0.4	0.3	0.2	0.1	0.1
15 or 85	3.6	2.9	2.5	2.1	1.6	1.1	0.9	0.8	0.5	0.4	0.3	0.2	0.1
20 or 80	4.0	3.3	2.8	2.3	1.8	1.3	1.0	0.9	0.6	0.4	0.3	0.2	0.1
25 or 75	4.3	3.5	3.1	2.5	1.9	1.4	1.1	1.0	0.6	0.4	0.3	0.2	0.1
30 or 70	4.6	3.7	3.2	2.6	2.0	1.4	1.2	1.0	0.6	0.5	0.3	0.2	0.1
35 or 65	4.8	3.9	3.4	2.8	2.1	1.5	1.2	1.1	0.7	0.5	0.3	0.2	0.2
50	5.0	4.1	3.5	2.9	2.2	1.6	1.3	1.1	0.7	0.5	0.4	0.2	0.2

1/ For a percentage and/or base of percentage not shown in the table, the formula given below may be used to calculate the standard error.

$$Se\ (\hat{p}) = \sqrt{\frac{5}{B}\hat{p}(100-\hat{p})}$$

B = Base of estimated percentage

$\hat{p}$ = Estimated percentage

Table L. Standard Error Adjustment Factors

Characteristic	Factor
EUROPEAN (excluding Spaniard):	
Single Ancestry--U.S. Total, Regions, Divisions, and All States	
English, German, Irish, Italian	1.5
Other European (excluding Spaniard) ancestry groups	1.2
At Least One or Multiple Ancestry--U.S. Total, Regions, Divisions and All States	1.6
NORTH AFRICAN AND MIDDLE EASTERNER (including South African*):	
Single Ancestry--U.S. Total, Regions, Divisions, and All States	1.2
At Least One or Multiple Ancestry--U.S. Total, Regions, Divisions, and All States	1.6
SUBSAHARA AFRICAN (excluding South African*):	
Single, Multiple, or At Least One Ancestry--U.S. Total, Regions, Divisions and the States of Alabama, Arkansas, District of Columbia, Florida, Georgia, Kentucky Louisiana, Maryland, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, West Virginia	2.3
All Other States	1.8
ASIAN (excluding Japanese and Middle Easterner):	
Single, Multiple, or At Least One Ancestry--U.S. Total, Regions, Divisions and All States	1.6
NON-SPANISH CARIBBEAN, CENTRAL, AND SOUTH AMERICAN:	
Single, Multiple, or At Least One Ancestry--U.S. Total, Regions, Divisions and All States	1.8
SPANISH:	
Mexican: Single, Multiple, or At Least One Ancestry--U.S. Total Regions, Divisions, and the States of Arizona, California, New Mexico, Oklahoma, Texas	2.1
All Other States	1.5
Puerto Rican: Single, Multiple, or At Least One Ancestry--U.S. Total, Regions, Divisions and the States of Massachussetts, New Jersey, New York	1.9
All Other States	1.5
Other Spanish Ancestry Groups: Single, Multiple, or At Least One Ancestry--U.S. Total, Regions, Divisions, and All States	1.5
PACIFIC (including Japanese):	
Single, Multiple, or At Least One Ancestry--U.S. Total, Regions, Divisions and the States of California and Hawaii	1.7
All Other States	1.2
NORTH AMERICAN:	
American Indian: Single, Multiple, or At Least One Ancestry--U.S. Total, Regions, Divisions, and the States of Alaska, Arizona, California, Colorado, Idaho, Minnesota, Montana, Nevada, New Mexico, Oklahoma, Oregon, South Dakota, Washington, Wyoming	1.7
All Other States	1.4
Aleut and Eskimo: Single, Multiple, or At Least One Ancestry--U.S. Total, Regions, Divisions, and the State of Alaska	1.6
All Other States	1.2
Canadian: Single, Multiple, or At Least One Ancestry--U.S. Total, Regions, Divisions, and All States	1.2
French Canadian: Single, Multiple, or At Least One Ancestry--U.S. Total, Regions, Divisions, and All States	1.2
Other North American Ancestry Groups: Single, Multiple, or At Least One Ancestry--U.S. Total, Regions, Divisions, and All States	1.2

*This category represents a general type of response, which may encompass several ancestry groups.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1983-381-956:1113